



UNITED STATES DEPARTMENT OF COMMERCE  
U.S. Census Bureau  
Washington, DC 20233-0001

November 22, 2021

MEMORANDUM FOR Carolyn M. Pickering  
Survey Director, Survey of Income and Program Participation

From: Anthony G. Tersine, Jr. *Anthony G. Tersine, Jr.*  
Chief, Demographic Statistical Methods Division

Subject: Source and Accuracy Statement for The Survey of Income and  
Program Participation Calendar Year 2020 Data Collection Public  
Use Files (S&A-25)

This memorandum includes the Source and Accuracy Statement for the Survey of Income and Program Participation (SIPP) data collected during Calendar Year 2020.

The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY22-POP001-0024

If you have any questions about this document, please contact Tracy Mattingly at [tracy.l.mattingly@census.gov](mailto:tracy.l.mattingly@census.gov) or 301-763-6445, Mahdi Sundukchi at [mahdi.s.sundukchi@census.gov](mailto:mahdi.s.sundukchi@census.gov) or 301-763-4228, Faith Nwaoha-Brown at [faith.n.nwaoha.brown@census.gov](mailto:faith.n.nwaoha.brown@census.gov) or 301-763-4696 or Jacob Eliason at [jacob.z.eliason@census.gov](mailto:jacob.z.eliason@census.gov) or 301-763-3331 .

cc:

Elizabeth Sinclair	(ADDP)	Jonathan Rothbaum
Kevin Tolliver		Bethany DeSalvo
David Waddington	(SEHSD)	Laryssa Mykyta
Jason Fields		Kurt Bauman
Hyon Shin		Sharon Stern
Adam Smith		Rachel Shattuck
Holly Fee		James Farber (DSMD)
Trudi Renwick		Tracy Mattingly
Stephanie Galvin		Mahdi Sundukchi
Alfred Gottschalck		Ashley Westra
Rose Kreider		Faith Nwaoha-Brown
Mark Klee		Julia Yang
Brian McKenzie		Brice Gnahore
Liana Fox		Jacob Eliason
Lindsay Monte		

*[This page is intentionally left blank]*

**Demographic Statistical Methods Division  
Sample Design and Estimation**

# **Source and Accuracy Statement for Calendar Year 2020 Data Collection of the Survey of Income and Program Participation (SIPP)**

**Version 1.0  
November 22, 2021**

**Faith Nwaoha-Brown  
Jacob Eliason  
Mahdi Sundukchi**

**James Farber, ADC  
Tracy Mattingly, Lead Scientist**



## Table of Contents

1.	Data Collection and Estimation.....	1
1.1	Source of Data .....	1
1.2	Weights Produced .....	5
1.3	Estimation.....	8
1.4	Population Controls .....	9
1.5	Use of Weights .....	10
2.	Accuracy of Estimates.....	14
2.1	Nonsampling Error.....	14
3.	Sampling Variability and Computation of Standard Errors .....	19
3.1	Confidence Intervals .....	19
3.2	Hypothesis Testing.....	19
3.3	Calculating Standard Errors for SIPP Estimates.....	21
3.4	Standard Error Parameters and Tables and Their Use .....	22
3.5	Standard Errors of Estimated Numbers .....	24
3.6	Standard Error of a Mean .....	24
3.7	Standard Error of an Aggregate .....	28
3.8	Standard Errors of Estimated Percentages .....	28
3.9	Standard Error of a Difference.....	30
3.10	Standard Error of a Median .....	31
3.11	Standard Errors of Ratios of Means and Medians .....	33
3.12	Standard Errors Using Software Packages .....	34
4.	References.....	36
5.	Generalized Variance Parameters and Tables .....	38

## List of Tables

Table 1. SIPP 2018 and SIPP 2020 Panels Household Counts, Sample Loss, and Weighted Response Rates .....	4
Table 2. SIPP 2018 and SIPP 2020 Panels Components of Type A Nonresponse by Wave .....	4
Table 3. Calendar Year Weights and Longitudinal Weights for CY18 through CY21 data in the Overlapping Panels Design.....	6
Table 4. CY20 Longitudinal Person Response Rates .....	7
Table 5. Coverage Ratios for December 2019 for Calendar Year Weights by Age, Race, Sex, and Hispanic Origin .....	16
Table 6. Coverage Ratios for December 2019 for Two-Year Longitudinal Weights ( <i>CY20 FINYR2</i> ) Weights by Age, Race, Sex, and Hispanic Origin.....	17
Table 7. Coverage Ratios for December 2019 for Three-Year Longitudinal Weights ( <i>CY20 FINYR3</i> ) Weights by Age, Race, Sex, and Hispanic Origin.....	18
Table 8. Calendar year 2019 SIPP key estimates computed from CY20 data. ....	21
Table 9. Cross-Sectional Generalized Variance Parameters for CY20 data .....	38
Table 10. Longitudinal Generalized Variance Parameters for CY20 FINYR2 .....	39
Table 11. Longitudinal Generalized Variance Parameters for CY20 FINYR3 .....	40

*[This page is intentionally left blank]*

## 1. Data Collection and Estimation

### 1.1 Source of Data

The data were collected in the 2018 and 2020 Panels of the Survey of Income and Program Participation (SIPP) <sup>1</sup>. The SIPP population universe is the civilian noninstitutionalized population living in the United States. The institutionalized population, which is composed primarily of the persons in correctional institutions and nursing homes (94 percent of the four million institutionalized people in the 2010 Census), is excluded from the SIPP universe.

Sample from each of the SIPP 2018 and SIPP 2020 Panels are located in 686 Primary Sampling Units (PSUs), each consisting of a county or a group of contiguous counties. Of these 686 PSUs, 252 are self-representing (SR) and 434 are non-self-representing (NSR). SR PSUs have a probability of selection of one and NSR PSUs have a probability of selection less than one. Within PSUs, Housing Units<sup>2</sup> (HUs) were systematically selected from the Master Address File (MAF), which is the U.S. Census Bureau's official inventory of known housing units. The MAF was created using the decennial censuses, as well as the U.S. Postal Service's Delivery Sequence File (DSF). The Census Bureau continues to update the MAF using the DSF and various automated, clerical, and field operations.

Housing Units were classified into two strata<sup>3</sup>, such that one stratum had a higher concentration of low income households than the other. We oversampled the low income stratum by 29 percent to increase the accuracy of estimates for low income households and program participation. Analysts are strongly encouraged to use the SIPP weights when computing estimates since households are not selected with equal probability.

Households in each SIPP Panel sample were scheduled to be interviewed at yearly intervals over a period of roughly four years beginning in February 2018 for the SIPP 2018 Panel and February 2020 for the SIPP 2020 Panel. The reference period for the interview questions is the preceding twelve-month calendar year. The most recent month, December, is designated reference month 12 and the earliest month, January, is reference month one. One cycle of interviewing covering the entire sample of a SIPP panel, using the same questionnaire, is called a wave. Data for up to 12 reference months are available for respondents and specific months available depend on a person's sample entry and/or exit date. During interviewing, Field Representatives (FRs) identify a reference person who is usually the owner or renter of the residence. A household's reference person may change from wave to wave due to changes in the household composition over time.

---

<sup>1</sup> This source and accuracy statement can also be accessed through the U.S. Census Bureau website at <http://www.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements.html>

<sup>2</sup> The SIPP selects housing units which may or may not be occupied at the time of interview; occupied housing units are referred to as households.

<sup>3</sup> Household income used to determine the low income and non-low income strata was obtained from the most recent American Community Survey (ACS) data that was available when the SIPP 2018 Panel and SIPP 2020 Panel samples were selected in 2017 and 2019 respectively.

Following the first interview of the 2018 Panel, the SIPP adopted an overlapping panels design in which a new panel is fielded annually and overlaps with one or more ongoing previous panel(s). Figure 1 depicts the overlapping panels design for SIPP 2018 and subsequent panels. Beginning in 2019 with the SIPP 2018 and SIPP 2019 Panels, sampled households from multiple panels are interviewed concurrently during each annual interview cycle and the combined data, covering the same reference period, is referred to by the year of interview. The *SIPP Calendar Year 2019 data collection* consisted of data from the SIPP 2018 Wave 2 and SIPP 2019 Wave 1 respondents who were interviewed in 2019 and covered the reference period from January to December 2018. The SIPP 2019 Panel was discontinued after the first wave due to low response. As a result, the *SIPP Calendar Year 2020 data collection* – hereafter referred to as *CY20 data*<sup>4</sup> – consists of data from the SIPP 2018 Wave 3 and SIPP 2020 Wave 1 respondents who were interviewed in 2020 and covers the reference period from January to December 2019.

The overlapping panel design was also implemented in early SIPP panels prior to the SIPP 1996 Panel. It allows analysts to combine data from multiple panels covering the same reference period for cross-sectional and longitudinal analyses, thus increasing the sample size, and decreasing standard errors SIPP key estimates (U.S. Census Bureau, 2020). Data Users should use data from all panels covering the same reference periods to compute SIPP key estimates in the overlapping panel design. Section 1.5 of the *2020 SIPP Users' Guide* further details the SIPP overlapping panel design, including comparisons to previous SIPP Panels.

Figure 1: Illustration of the Overlapping Panels Design for the SIPP 2018 and Subsequent Panels.

Year of interview										
	2017	2018	2019	2020	2021	2022	2023	2024	2025	...
Panel		2018 Panel								
			2019 Panel <sup>1</sup>							
				2020 Panel						
					2021 Panel					
						2022 Panel				
							2023 Panel			
								2024 Panel		

<sup>1</sup> The SIPP 2019 Panel was terminated after Wave 1 due to low response.

The SIPP 2018 Panel began with a sample of about 53,500 HUs in Wave 1. About 8,800 of these HUs were found to be vacant, demolished, converted to nonresidential use, or otherwise ineligible for the survey. FRs were able to obtain interviews from about 26,000 of the eligible households. FRs were unable to interview approximately 18,500 eligible households in the panel because the occupants: (1) refused to be interviewed; (2) could not be found at home; (3) were temporarily absent; or (4) were otherwise unavailable. Thus, occupants of about 58 percent of all eligible households participated in the first interview of the 2018 Panel.

<sup>4</sup> CY20 data is referred to as the 2020 SIPP in the 2020 Users' Guide.



The SIPP 2020 Panel began with a sample of about 26,000 HUs in Wave 1. About 4,000 of these HUs were found to be vacant, demolished, converted to nonresidential use, or otherwise ineligible for the survey. FRs were able to obtain interviews from about 8,000 of the eligible households. FRs were unable to interview approximately 14,000 eligible households in the panel because the occupants: (1) refused to be interviewed; (2) could not be found at home; (3) were temporarily absent; or (4) were otherwise unavailable. Thus, occupants of about 36 percent of all eligible households participated in the first interview of the 2020 Panel.

Only original sample people (interviewed persons in Wave 1 sample households) and people living with them are eligible to be interviewed in subsequent waves. Original sample people who move from their Wave 1 address to a new address in later waves are still included in the SIPP sample. However, FRs attempt telephone interviews in lieu of in-person interviews if their new address is more than 100 miles from a SIPP sample area. Approximately 14,000 of the 27,000 SIPP 2018 Panel eligible HUs were interviewed in Wave 3. Therefore, the CY20 data has a total of approximately 22,000 interviews out of 49,000 eligible HUs, resulting in a cross-sectional weighted response rate of 45 percent.

Since the SIPP follows all original sample members, those that form new households are also included in the SIPP sample. This expansion of original households can be estimated within the interviewed sample but is impossible to determine within the noninterviewed sample. Therefore, a growth factor based on the growth in the known sample is used to estimate the unknown expansion of the noninterviewed households. Growth factors account for the additional nonresponse stemming from the expansion of noninterviewed households and are calculated for Wave 2 and later waves. They provide a more accurate estimate of the weighted counts of noninterviewed households at each wave.

There are two categories of noninterviewed households: Type A and Type D. Type A noninterviewed households are eligible households where the interviewer obtains no interview. Type D noninterviewed households are previously interviewed households who moved to an unknown address or moved more than 100 miles from a SIPP interviewer and no telephone interview could be conducted. As a result, Type D noninterviews only occur from Wave 2 onwards. To calculate this loss of sample, or “sample loss,” we use Formula (1):

$$\text{Sample Loss} = \frac{(A_1 \times GF_c) + A_c + D_c}{I_c + (A_1 \times GF_c) + A_c + D_c} \quad (1)$$

where:

$A_1$  = weighted number of Type A noninterviewed households in Wave 1

$A_c$  = weighted number of Type A noninterviewed households in the current wave

$D_c$  = weighted number of Type D noninterviewed households in the current wave

$I_c$  = weighted number of interviewed households in the current wave

$GF_c$  = growth factor associated with the current wave.

Weighted sample loss for each wave of the SIPP 2018 and 2020 Panels calculated using equation (1) are tabulated in Table 1. The distribution of Type A nonrespondents by nonresponse reason is shown in Table 2.

Table 1. SIPP 2018 and SIPP 2020 Panels Household Counts, Sample Loss, and Weighted Response Rates

	Eligible Households <sup>1</sup>	Interviewed Households	Type A Households		Type D Households		Growth Factor	Cumulative Weighted Response Rates (percent)	Weighted Sample Loss (percent)
			Total	Weighted Rate (percent)	Total	Weighted Rate (percent)			
	SIPP 2018 Panel								
Wave 1	45,000	26,000	18,500	41.56	–	–	–	58.44	41.56
Wave 2	24,000	14,500	8,900	36.78	650	2.47	1.00	34.33	65.67
Wave 3	27,000	14,000	12,500	45.97	500	1.74	1.02	30.55	69.45
	SIPP 2020 Panel								
Wave 1	22,000	8,000	14,000	63.62	–	–	–	36.38	63.62

Source: U.S. Census Bureau, Survey of Income and Program Participation Calendar Year 2018, Calendar Year 2019, and Calendar Year 2020.

<sup>1</sup> Interviewed and noninterviewed households may not sum up to eligible households due to rounding.

Table 2. SIPP 2018 and SIPP 2020 Panels Components of Type A Nonresponse by Wave

Panel	Wave	Language Problem	Unable to Locate	No One Home	Temporarily Absent	Household Refused	Other
<b>SIPP 2018</b>	<b>1</b>	1.46	0.21	13.29	2.24	76.14	6.65
	<b>2</b>	0.81	0.00	11.50	1.72	78.12	7.86
	<b>3</b>	0.59	0.00	13.98	0.20	38.70	46.54
<b>SIPP 2020</b>	<b>1</b>	0.00	0.75	22.54	0.12	25.35	51.23

Source: U.S. Census Bureau, Survey of Income and Program Participation Calendar Year 2018, Calendar Year 2019, and Calendar Year 2020.

## 1.2 Weights Produced

The SIPP produces three weights for cross-sectional and longitudinal analyses: monthly weights (*WPFINWGT*), calendar year weights<sup>5</sup> (*WPFINWGT* when *MONTHCODE*=12), and longitudinal weights<sup>6</sup> (*FINYR*<*n*>, where *n* is the number of calendar years in the reference period covered by a longitudinal weight). Monthly and calendar year weights are cross-sectional weights because they are computed using data obtained from a single interview. Longitudinal weights, however, are computed by combining data from multiple interviews – i.e., two or more waves.

Monthly weights are used to calculate estimates for each of the 12 months within a calendar year. For each reference month, respondents who were in the SIPP survey universe and for whom data were obtained, receive positive monthly weights.

Calendar year weights cover the reference period from January to December of a specified calendar year and can be used to calculate estimates for any period within the year. Each CYyy data in the overlapping panels design covers the reference period from January to December 20yy-1, where 20yy is the year of interview and 20yy-1 is the reference calendar year preceding the year of interview. Therefore, CY20 data calendar year weights can be used to compute monthly<sup>7</sup>, quarterly, and annual estimates for any time between January and December 2019.

SIPP calendar year weights are based on the SIPP survey universe in December of a designated year. All interviewed persons in the SIPP survey universe who have positive December monthly weights are assigned calendar year weights equal to their monthly December weights. As a result, separate calendar year weight and replicate weight files are not produced for the CY20 data. Table 3 shows the reference period for calendar and longitudinal weights for the CY18 through CY21<sup>8</sup> data releases. The calendar year weights for years 2018, 2019, and 2020 incorporate data from the SIPP 2018 Panel through SIPP 2021 Panel in accordance with the overlapping panels design described in Section 1.1 and Figure 1.

Longitudinal weights cover reference periods of two or more calendar years and can be used to calculate estimates in this interval. The eligible sample cohort for longitudinal weights are persons who had positive calendar year weights in the first year of the longitudinal reference period. Eligible persons are classified as *interviewed for longitudinal weights* if they have positive calendar year weights in all subsequent years in the longitudinal reference period and

---

<sup>5</sup> Calendar year weights in the SIPP 2014 Panel Wave 1 through Wave 4 and CY18 source and accuracy documents were referred to as CY<year> where year was the calendar year reference period covered by the weights.

<sup>6</sup> Longitudinal weights were also referred to as Panel weights in the SIPP 1996 through SIPP 2014 Panels because the weights were computed from sample persons in a single panel. Longitudinal weights in the overlapping panels design however can include sample persons from more than one panel as depicted in Table 3 CY20 FINYR2 weights.

<sup>7</sup> We recommend analysts use monthly weights instead of calendar year weights to compute monthly estimates as using calendar year weights may result in fewer sample persons.

<sup>8</sup> Table 3 shows calendar year weights and longitudinal weights from CY18 through CY21 data releases because the first SIPP Panel in the overlapping panels design (SIPP 2018 Panel) ends in CY21 data release.

receive positive longitudinal weights<sup>9</sup>. Eligible persons who do not have calendar year weights in at least one year of the reference period are categorized as *noninterviewed for longitudinal weights* and are not assigned longitudinal weights.

Beginning with CY19 data, each data release under the overlapping panels design will include one, two, or three longitudinal weights – *FINYR2*, *FINYR3*, and *FINYR4* – covering a reference period of two, three and four earlier calendar years respectively. Each CYyy *FINYRn* longitudinal weight covers an *n*-year reference period ending in year 20yy-1 with the last interview occurring in year 20yy. The number of longitudinal weights published with each CYyy data release depends on the length of each ongoing panel whose sampled households were interviewed in 20yy. Table 3 lists the longitudinal weights published with CY18 through CY21 data files.

Table 3. Calendar Year Weights and Longitudinal Weights for CY18 through CY21 data in the Overlapping Panels Design.

SIPP Panel	CYyy data, where 20yy is the year of interview			
	CY18	CY19	CY20	CY21
2018	Wave 1	Wave 2	Wave 3	Wave 4
2019		Wave 1	SIPP 2019 Panel Terminated	
2020			Wave 1	Wave 2
2021				Wave 1
Cross-Sectional Reference Calendar Years Covered by CYyy File				
	2017	2018	2019	2020
Calendar Year Weights Produced for CYyy File <sup>1</sup>				
	WPFINWGT for MONTHCODE=12	WPFINWGT for MONTHCODE=12 <sup>1</sup>	WPFINWGT for MONTHCODE=12 <sup>2</sup>	WPFINWGT for MONTHCODE=12 <sup>3</sup>
Longitudinal Weights Produced for CYyy File				
FINYR2	–	✓	✓	✓
FINYR3	–	–	✓	✓
FINYR4	–	–	–	✓

<sup>1</sup> 2018 calendar year weights incorporate data from SIPP 2018 Panel Wave 2 and SIPP 2019 Panel Wave 1 sample.

<sup>2</sup> 2019 calendar year weights incorporate data from SIPP 2018 Panel Wave 3 and SIPP 2020 Panel Wave 1 sample.

<sup>3</sup> 2020 calendar year weights incorporate data from SIPP 2018 Panel Wave 4, SIPP 2020 Panel Wave 2, and SIPP 2021 Panel Wave 1 sample.

CY20 data includes two longitudinal weights, *FINYR2* and *FINYR3*. *FINYR2* weights cover the two-year reference period from January 2018 to December 2019. The eligible sample cohort for

<sup>9</sup> Some eligible persons – including children in households without any adults eligible for longitudinal weights, and persons listed in error in prior waves – are excluded from the longitudinal weighting procedure. As a result, the total number of eligible persons for CY19 *FINYR2* and CY20 *FINYR3* longitudinal weights differ though both weights have the same eligible sample cohort – persons with positive December 2017 weights.

CY20 *FINYR2* weights are persons who had positive calendar year weights in the first year of the reference period, i.e., SIPP 2018 Panel Wave 2 sample persons with positive December 2018 weights. About 35,000 persons were eligible for CY20 *FINYR2* weights and 22,000 of them were classified as interviewed and received positive longitudinal weights for a weighted response rate of 63 percent. Longitudinal response is computed at the person level as household compositions may change over the longitudinal reference period. Table 4 shows the response rate for each CY20 longitudinal weight.

*FINYR3* weights cover the three-year reference period from January 2017 to December 2019. The eligible sample cohort for CY20 *FINYR3* weights are persons who had positive calendar year weights in the first year of the reference period, i.e., SIPP 2018 Wave 1 sample persons with positive December 2017 weights. About 63,500 persons were eligible for CY20 *FINYR3* weights and 21,000 of them were classified as interviewed and received positive longitudinal weights for a weighted response rate of 33 percent.

Table 4. CY20 Longitudinal Person Response Rates

Longitudinal Weight	Eligible Persons	Interviewed Persons	Noninterviewed Persons	Weighted Response Rate (percent)
FINYR2	35,000	22,000	13,000	63.4
FINYR3	63,500	21,000	42,500	33.1

Source: U.S. Census Bureau, Survey of Income and Program Participation Calendar Year 2018, Calendar Year 2019, and Calendar Year 2020.

Unlike the SIPP 2014 and earlier SIPP Panels, longitudinal weights in the overlapping panel design may include persons from multiple panels. The eligible sample cohort for CY21 *FINYR2* weights is highlighted in Table 3 and consists of SIPP 2018 Panel Wave 3 and SIPP 2020 Panel Wave 1 sample persons who had positive calendar year weights in 2019.

### 1.3 Estimation

The SIPP estimation procedure involves several stages of weight adjustments to derive the final person level weights. The cross-sectional weighting procedure for overlapping panels data – pooled data from two or more panels whose sample persons were interviewed in the same year and hence have the same reference period as shown in Figure 1 and Table 3 – is divided into two steps: (1) noninterview adjustment and (2) second stage adjustment.

Household noninterview adjustment is conducted separately for each panel due to the different selection probabilities and response rates across panels. The nonresponse adjustment procedure depends on the specific waves in each panel. CY20 household noninterview adjustments adjust for nonresponse among SIPP 2018 Panel Wave 3 and SIPP 2020 Panel Wave 1 sample households. For Wave 1 households, each eligible household is first given a base weight ( $BW$ ) equal to the inverse of its probability of selection. Base weights are multiplied by a weighting control factor ( $WCF$ ) to adjust for subsampling done in the field when the number of sampled units is much larger than expected. A noninterview adjustment factor ( $F_{N1}$ ) is then applied to account for eligible households that FRs could not interview to create the Wave 1 noninterview adjusted weights and assigned to all household members.

For Wave 3 households, each eligible household is assigned an initial weight equal to its Wave 1 household noninterview adjusted weight ( $INITwgt_{w1}$ ). Initial weights are multiplied by movers' adjustment factors ( $MF$ ) to account for multiple chances of selection of movers<sup>10</sup>. The resulting movers' weights are multiplied by the noninterview adjustment factor ( $F_{N2}$ ) to adjust for household nonresponse in Wave 3 and create household noninterview adjusted weights, which are assigned to all household members.

Next, the second stage adjustment is implemented on the pooled sample from all contributing panels. Noninterview adjusted weights are multiplied by an *overlapping panel factor* ( $OPF_{panel}$ ) to adjust for combining sample households from multiple panels and a second stage adjustment factor ( $F_{2S}$ ) to create final person weights. The second stage adjustment equalizes married spouses' weights and rakes the sum of person weights to independent population controls (benchmark population estimates) by age, race, sex, Hispanic origin, and state of residence. Additional details of the weighting process are in Tersine (2021a) and Tersine (2021b).

In summary, the final cross-sectional weight ( $FW_c$ ) for each month  $c$ , in CY20 data is

$$\begin{aligned} & BW * WCF * F_{N1} * OPF_{s20} * F_{2S} && \text{for SIPP 2020 Panel sample persons} \\ & INITwgt_{w1} * MF * F_{N2} * OPF_{s18} * F_{2S} && \text{for SIPP 2018 Panel sample persons} \end{aligned}$$

<sup>10</sup> Movers – persons who move into SIPP sample households after Wave 1 interviews – have two chances to become SIPP sample persons: (a) selection into original SIPP sample households in Wave 1 or (b) selection by moving into a sample household after Wave 1.

For each longitudinal weight, FINYR#, eligible persons are given an initial weight ( $IW_{\#}$ ) equal to their cross-sectional household noninterview adjusted weight for the first year of the reference period. A noninterview adjustment factor ( $FP_{ni\#}$ ) is then applied to account for person level nonresponse. A Longitudinal Overlapping Panel Factor ( $LOPF_{panel}$ ) is applied to the noninterview adjusted weights if the longitudinal cohort consists of persons from more than one panel. Finally, a second stage adjustment ( $FP_{2S\#}$ ) is applied to adjust the noninterview weights to independent population controls for December 20yy-1, the last month of the longitudinal reference period. Spouse weights are not equalized in the longitudinal weighting procedure.

The final longitudinal weight is

$$\begin{aligned} IW_3 * FP_{ni3} * LOPF_{2018} * FP_{2S3} & \text{ for FINYR2} \\ IW_2 * FP_{ni2} * LOPF_{2018} * FP_{2S2} & \text{ for FINYR3} \end{aligned}$$

with  $LOPF_{2018} = 1$  since only SIPP 2018 Panel sample persons are eligible for CY20 FINYR2 and FINYR3 weights. Additional details of the weighting process are in Tersine (2021a) and Tersine (2021b).

#### 1.4 Population Controls

The SIPP estimation procedure adjusts weighted sample estimates to agree with independently derived population estimates of the civilian noninstitutionalized population. This attempts to correct for undercoverage and thereby reduces the mean square error of the estimate. The national and state level population controls are obtained directly from the Population Division and are prepared each month to agree with the most current set of population estimates released by the Census Bureau's population estimates and projections program.

Longitudinal weights in the overlapping panels design use the population controls for the last month of the reference period. As a result, all longitudinal weights and calendar year weights released with each CYyy data sum up to the population controls for December 20yy-1.

The national level controls are distributed by demographic characteristics as follows:

- Age, Sex, and Race (White alone, Black alone, and all other groups combined)
- Age, Sex, and Hispanic Origin

The state level controls are distributed by demographic characteristics as follows:

- State, Age, and Sex
- State, Hispanic origin
- State, Race (Black alone, all other groups combined)

The estimates begin with the latest decennial census as the base and incorporate the latest available information on births and deaths along with the latest estimates of net international migration. The net international migration component in the population estimates includes a combination of:

- Legal migration to the U.S.,
- Emigration of foreign born and native people from the U.S.,
- Net movement between the U.S. and Puerto Rico,
- Estimates of temporary migration, and
- Estimates of net residual foreign-born population, which include unauthorized migration.

Because the latest available information on these components lags the survey date, it is necessary to make short-term projections of these components to develop the estimates for the survey date.

### 1.5 Use of Weights

The SIPP monthly, calendar year, and longitudinal weights are produced at the person level and are intended for analyzing data at the person level. Every interviewed person in the SIPP universe for a given reference month – and for whom data were obtained – has a person month weight for that reference month. Likewise, persons who had data for the month of December have calendar year weights, and persons categorized as *interviewed for a longitudinal weight* are assigned corresponding longitudinal weights. Chapter 7 of the 2020 *SIPP Users' Guide* provides additional information on how to use the survey weights.

The SIPP 2008 and earlier SIPP Panels' public use files also contained household, family, and related subfamily monthly weights for analyzing the data at the appropriate household and family levels. These weights were set to be the person month weight of the household, family, or subfamily reference person for that reference month<sup>11</sup>. For the SIPP 2014 Panel and subsequent panels, the household structure of an interviewed unit is only set for the interview month. Up to five addresses are recorded for each person for the reference period, so interviewed persons can live in different households depending on the reference month. Therefore, for each reference month it is possible to tell which interviewed persons lived together and their relationships to each other, but the files do not specify a household ID or reference person for each of the reference months. The same is true for families. If a data user would like to conduct analysis at the household or family level, person weights can be used to specify a single household or family weight. One option is to use the average of the person month weights for all persons in the household or family. Another option is to specify a household or family reference person and use his or her person month weight as the household or family weight.

---

<sup>11</sup> Only person-level calendar year and longitudinal weights were available in previous SIPP panels.



All estimates may be divided into two broad categories: longitudinal and cross-sectional. Longitudinal estimates require that data records for each person be linked across interviews, whereas cross-sectional estimates do not. For example, estimating the average duration of unemployment from October 2018 to September 2019 requires linking records from CY19 data and CY20 data, restricting to only the sample persons who were interviewed in both years, and would be a longitudinal estimate. Cross-sectional estimates can combine data from different interviews only at the aggregate level because there is no linkage between interviews; for example, comparing the national unemployment rate in September 2018 to that in September 2019.

*Data users are strongly encouraged to use the weights provided as instructed. Weights can vary due to the sample design, the level of nonresponse, and the time period of the estimate. Calculating estimates without using the weights as instructed will produce erroneous results.*

Some basic types of estimates that can be constructed using monthly, calendar year, and longitudinal weights are described below in terms of estimated numbers. More complex estimates, such as percentages, averages, ratios, etc., can be constructed from the estimated numbers.

1. The number of people who have ever experienced a characteristic during a given time period.

To construct such an estimate, use the person weight for the shortest time period which covers the entire time period of interest. Then sum the weights over all people who possessed the characteristic of interest at some point during the time period of interest. For example, to estimate the number of people who received Supplemental Nutrition Assistance Program (SNAP) benefits in January 2019, sum the monthly weights (*WPFINWGT*, with *MONTHCODE*=1) of all persons who received SNAP benefits in January 2019.

To estimate the number of persons who experienced poverty at any point between January 2018 and December 2019, sum the two-year longitudinal weights (*FINYR2*) of all persons who were in poverty during this time frame. The three-year longitudinal weights (*FINYR3*) can also be used in computing this estimate but may result in fewer sample persons and larger standard errors.

To estimate the number of persons who received any unemployment income between 2017 and 2019, sum the three-year longitudinal weights (*FINYR3*) of all persons who received unemployment income anytime between January 2017 and December 2019.

2. The amount of a characteristic accumulated by people during a given time period.

To construct such an estimate, use the person weight for the shortest time period which covers the entire time period of interest. Compute the product of the weight and the

amount of the characteristic and sum this product over all appropriate people. For example, to estimate the aggregate 2019 annual income of people who were employed during all 12 months of the year, multiply the CY20 calendar year weights (*WPFINWGT*, with *MONTHCODE=12*) of all persons who were employed in all months of 2019 by their annual income and sum the resulting products.

3. The average number of consecutive months of possession of a characteristic (i.e., the average spell length for a characteristic) during a given time period.

For example, one could estimate the average length of each spell of Supplemental Security Income (SSI) receipt during 2019. One could also estimate the average spell of unemployment that elapsed before a person found a new job. To construct such an estimate, first identify the people who possessed the characteristic at some point during the period of interest. Then create two sums of these persons' appropriate weights: (1) sum the product of the weight times the number of months the spell lasted and (2) sum the weights only. The estimated average spell length in months is computed as (1) divided by (2). A person who experienced two spells during the time period of interest would be treated as two people and appears twice in sums (1) and (2). An alternate method of calculating the average can be found in the Section 3.6.

4. The number of month-to-month changes in the status of a characteristic (i.e., number of transitions) summed over every set of two consecutive months during the time period of interest.

To construct such an estimate, sum the appropriate person weights each time a change is reported between two consecutive months during the period of interest. For example, to estimate the number of people who were in poverty in July 2019 and transitioned out of poverty in August 2019, add the CY20 calendar year weights of each person who had such a change.

To estimate the number of changes in monthly earned income during the 2019 fiscal year (September 2018 to September 2019) use CY20 FINYR2 weights. Sum the weighted number of people who had a change between September and October, between October and November, between November and December, ..., and between August and September. The three-year longitudinal weights (FINYR3) can also be used in computing this estimate but may result in fewer sample persons and hence larger standard errors.

Spell and transition estimates should be used with caution because of the biases that are associated with them. Sample people tend to report the same status of a characteristic for all months of a reference period. This tendency also affects transition estimates in that, for many characteristics, the number of characteristics, the number of month-to-month transitions reported between the last month of one reference period and the first month of the next reference period are much greater than the number of

reported transitions between any two months within a reference period. Additionally, spells extending before or after the time period of interest are cut off (censored) at the boundaries of the time period. If they are used in estimating average spell length, a downward bias will result.

5. Monthly estimates of a characteristic averaged over consecutive months.

For example, one could estimate the monthly average number of Temporary Assistance for Needy Families (TANF) recipients from July 2019 through December 2019. To construct such an estimate, first form an estimate for each month in the time period of interest. Sum the CY20 calendar year weight (i.e., *WPFINWGT*, with *MONTHCODE=12*), of all persons who possessed the characteristic of interest during each of the six months of interest. Then sum the monthly estimates and divide by six, the number of months.

## 2. Accuracy of Estimates

SIPP estimates are based on a sample; they may differ somewhat from the figures that would have been obtained from a complete census using the same questionnaire, instructions, and enumerators. There are two types of errors possible in an estimate based on a sample survey: sampling and nonsampling errors. For a given estimator, the difference between an estimate based on a sample and the estimate that would result if the sample were to include the entire population is known as sampling error. For a given estimator, the difference between the estimate that would result if the sample were to include the entire population and the true population value being estimated is known as nonsampling error. We can provide estimates of the magnitude of SIPP sampling error, but this is not true of nonsampling error.

### 2.1 Nonsampling Error

Nonsampling errors can be attributed to many sources:

- inability to obtain information about all cases in the sample
- definitional difficulties
- differences in the interpretation of questions
- inability or unwillingness on the part of the respondents to provide correct information
- errors made in the following: collection such as in recording or coding the data, processing the data, estimating values for missing data
- biases resulting from the differing recall periods caused by the interviewing pattern used and undercoverage.

Quality control and edit procedures were used to reduce errors made by respondents, coders, and interviewers. More detailed discussions of the existence and control of nonsampling errors in the SIPP can be found in the *SIPP Quality Profile, 1998 SIPP Working Paper Number 230*, issued June 1998 (Kalton, 1998), the *SIPP 2018 Users' Guide* (U.S. Census Bureau, 2020) and the *SIPP 2020 Users' Guide* (U.S. Census Bureau, 2021).

Undercoverage in SIPP results from missed HUs and missed persons within sample HUs. It is known that undercoverage varies with age, race, and sex. Generally, undercoverage is larger for males than for females and larger for Blacks than for non-Blacks. Ratio estimation to independent age-race-sex-Hispanic origin population controls during the second stage adjustment step of the SIPP weighting procedure partially corrects for the bias due to survey undercoverage. However, biases exist in the estimates to the extent that persons in missed households or missed persons in interviewed households have characteristics different from those of interviewed persons in the same age-race-sex-Hispanic origin group.

A common measure of survey coverage is the coverage ratio, the estimated population before second stage adjustment divided by the independent population control. Tables 5, 6, and 7 show coverage ratios for age-sex-race-Hispanic origin groups in December 2019, the control month for CY20 calendar year and longitudinal weights, using calendar year weights and two-year and three-year longitudinal weights respectively prior to post stratification ratio

adjustments. The SIPP coverage ratios exhibit some variability from month to month. Other Census Bureau household surveys (e.g. the Current Population Survey) experience similar coverage.

Caution should be exercised when comparing the SIPP CY20 data with data from other SIPP products<sup>12</sup> or with data from other surveys. The comparability problems are caused by sources such as the seasonal patterns for many characteristics, different nonsampling errors, and different concepts and procedures. Refer to the *SIPP Quality Profile* (Kalton, 1998) for known differences with data from other sources and further discussions.

---

<sup>12</sup> Analysts should be cautious when comparing estimates from the Re-engineered SIPP Panels (i.e., The SIPP 2014 and subsequent SIPP panels) to previous panels as the SIPP was redesigned following the SIPP 2008 Panel. The questionnaire, interview frequency, collection format, and sample size of the SIPP 2014 and later panels differ from those of previous SIPP panels. Detailed information on the reengineered SIPP are provided in the 2014 SIPP Users' Guide and the 2018 SIPP Users' Guide.

Table 5. Coverage Ratios for December 2019 for Calendar Year Weights by Age, Race, Sex, and Hispanic Origin

Age	All Persons		White non-Hispanic		Black non-Hispanic		Other non-Hispanic		Hispanic (of any race)	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
<b>&lt;15</b>	0.71	0.68	0.76	0.69	0.56	0.60	0.77	0.79	0.66	0.65
<b>15</b>	0.81	0.74	0.89	0.77	0.68	0.67	0.70	0.71	0.79	0.75
<b>16-17</b>	0.71	0.78	0.72	0.82	0.72	0.69	0.70	0.77	0.67	0.75
<b>18-19</b>	0.76	0.75	0.79	0.79	0.76	0.65	0.72	0.73	0.71	0.74
<b>20-21</b>	0.68	0.67	0.74	0.71	0.53	0.49	0.71	0.71	0.61	0.67
<b>22-24</b>	0.72	0.61	0.78	0.62	0.49	0.53	0.70	0.72	0.69	0.59
<b>25-29</b>	0.65	0.66	0.71	0.69	0.55	0.56	0.65	0.70	0.56	0.63
<b>30-34</b>	0.68	0.68	0.75	0.71	0.51	0.59	0.68	0.72	0.59	0.66
<b>35-39</b>	0.74	0.71	0.78	0.75	0.57	0.60	0.84	0.74	0.65	0.67
<b>40-44</b>	0.73	0.77	0.77	0.80	0.61	0.76	0.81	0.76	0.64	0.72
<b>45-49</b>	0.75	0.78	0.78	0.81	0.73	0.66	0.86	0.89	0.64	0.75
<b>50-54</b>	0.84	0.85	0.88	0.86	0.77	0.85	0.92	0.90	0.71	0.79
<b>55-59</b>	0.88	0.94	0.90	0.94	0.91	0.97	0.93	0.98	0.75	0.88
<b>60-61</b>	0.99	1.07	1.03	1.10	0.89	0.99	0.96	1.00	0.85	1.02
<b>62-64</b>	0.99	1.05	1.01	1.07	0.96	1.00	1.00	1.03	0.84	1.00
<b>65-69</b>	1.11	1.19	1.15	1.23	1.06	1.19	1.17	1.09	0.86	1.00
<b>70-74</b>	1.15	1.21	1.19	1.23	1.07	1.26	1.16	1.14	0.92	1.02
<b>75-79</b>	1.32	1.24	1.38	1.29	1.07	1.20	1.17	1.12	1.01	1.03
<b>80-84</b>	1.30	1.28	1.35	1.31	1.10	1.20	1.19	1.13	1.02	1.10
<b>85+</b>	1.28	1.20	1.33	1.22	1.02	1.21	1.17	1.06	1.01	1.05

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2020.

Table 6. Coverage Ratios for December 2019 for Two-Year Longitudinal Weights (CY20 FINYR2)  
Weights by Age, Race, Sex, and Hispanic Origin

Age	All Persons		White non-Hispanic		Black non-Hispanic		Other non-Hispanic		Hispanic (of any race)	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
<15	0.63	0.59	0.68	0.62	0.45	0.45	0.76	0.75	0.57	0.54
15	0.81	0.73	0.86	0.80	0.75	0.67	0.69	0.68	0.79	0.64
16-17	0.72	0.71	0.74	0.78	0.74	0.65	0.76	0.71	0.67	0.63
18-19	0.68	0.66	0.68	0.73	0.73	0.58	0.76	0.67	0.60	0.58
20-21	0.61	0.56	0.64	0.60	0.51	0.33	0.79	0.61	0.54	0.54
22-24	0.63	0.47	0.66	0.49	0.49	0.37	0.71	0.68	0.58	0.40
25-29	0.50	0.51	0.55	0.54	0.37	0.36	0.63	0.64	0.42	0.50
30-34	0.58	0.56	0.65	0.58	0.42	0.43	0.61	0.60	0.47	0.59
35-39	0.65	0.64	0.68	0.67	0.53	0.45	0.77	0.70	0.57	0.66
40-44	0.67	0.72	0.73	0.72	0.50	0.74	0.72	0.72	0.60	0.71
45-49	0.72	0.70	0.72	0.70	0.72	0.60	0.94	0.98	0.59	0.64
50-54	0.75	0.81	0.79	0.80	0.63	0.81	0.92	0.99	0.66	0.76
55-59	0.86	0.89	0.84	0.87	1.01	1.03	1.05	0.98	0.74	0.87
60-61	0.95	1.03	0.97	1.06	0.99	0.95	0.99	1.00	0.78	0.95
62-64	0.99	0.98	1.00	0.99	1.01	0.98	1.03	0.93	0.82	0.97
65-69	1.11	1.18	1.12	1.17	1.11	1.33	1.34	1.24	0.85	1.04
70-74	1.13	1.25	1.16	1.22	1.00	1.45	1.37	1.37	0.87	1.11
75-79	1.35	1.23	1.42	1.23	0.94	1.36	1.42	1.19	1.03	1.10
80-84	1.31	1.34	1.34	1.35	1.21	1.33	1.33	1.31	1.04	1.20
85+	1.38	1.24	1.41	1.24	1.06	1.42	1.36	1.30	1.10	1.08

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2019, and Calendar Year 2020.

Table 7. Coverage Ratios for December 2019 for Three-Year Longitudinal Weights (CY20 FINYR3)  
Weights by Age, Race, Sex, and Hispanic Origin

Age	All Persons		White non-Hispanic		Black non-Hispanic		Other non-Hispanic		Hispanic (of any race)	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
<15	0.69	0.66	0.75	0.70	0.50	0.49	0.82	0.78	0.62	0.61
15	0.93	0.83	0.97	0.89	0.83	0.83	0.83	0.79	0.93	0.72
16-17	0.83	0.82	0.84	0.87	0.86	0.79	0.88	0.83	0.78	0.72
18-19	0.81	0.79	0.81	0.86	0.85	0.73	0.92	0.79	0.73	0.68
20-21	0.69	0.64	0.72	0.69	0.55	0.44	0.88	0.69	0.65	0.60
22-24	0.70	0.54	0.73	0.55	0.52	0.50	0.84	0.80	0.64	0.44
25-29	0.55	0.56	0.59	0.56	0.39	0.45	0.69	0.67	0.49	0.58
30-34	0.65	0.62	0.72	0.63	0.47	0.48	0.68	0.65	0.56	0.68
35-39	0.75	0.72	0.77	0.73	0.57	0.52	0.89	0.84	0.71	0.78
40-44	0.78	0.84	0.82	0.81	0.62	0.92	0.84	0.87	0.74	0.87
45-49	0.85	0.81	0.84	0.79	0.88	0.73	1.10	1.11	0.72	0.78
50-54	0.88	0.94	0.91	0.91	0.71	0.93	1.05	1.12	0.81	0.92
55-59	1.02	1.05	0.98	0.99	1.24	1.29	1.22	1.19	0.91	1.07
60-61	1.13	1.21	1.15	1.23	1.21	1.15	1.17	1.17	0.99	1.17
62-64	1.16	1.17	1.17	1.16	1.22	1.22	1.18	1.12	1.00	1.22
65-69	1.31	1.38	1.31	1.35	1.34	1.60	1.61	1.47	1.05	1.28
70-74	1.47	1.44	1.51	1.40	1.25	1.77	1.63	1.60	1.22	1.36
75-79	1.44	1.48	1.47	1.46	1.14	1.62	1.68	1.42	1.15	1.42
80-84	1.50	1.51	1.52	1.50	1.45	1.61	1.59	1.54	1.26	1.40
85+	1.46	1.43	1.47	1.40	1.28	1.73	1.57	1.50	1.22	1.29

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2018, Calendar Year 2019, and Calendar Year 2020.



### 3. Sampling Variability and Computation of Standard Errors

Standard errors indicate the magnitude of the sampling error. They also partially measure the effect of some nonsampling errors in response and enumeration, but do not measure any systematic biases in the data. The standard errors for the most part measure the variations that occurred by chance because a sample rather than the entire population was surveyed.

#### 3.1 Confidence Intervals

The sample estimate and its standard error enable one to construct a confidence interval. A confidence interval is a range about a given estimate that has a known probability of including the result of a complete enumeration. For example, if all possible samples were selected, each of these being surveyed under essentially the same conditions and using the same sample design, and if an estimate and its standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate of all possible samples would include the actual value of the population parameter.
2. Approximately 90 percent of the intervals from 1.645 standard errors below the estimate to 1.645 standard errors above the estimate of all possible samples would include the actual value of the population parameter.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate of all possible samples would include the actual value of the population parameter.

The estimate derived from all possible samples may or may not be contained in any computed interval. However, for a particular sample, one can say with a specified confidence that the estimate derived from all possible samples is included in the confidence interval.

#### 3.2 Hypothesis Testing

Standard errors may also be used for hypothesis testing, a procedure for distinguishing between population characteristics using sample estimates. The most common types of hypotheses tested are 1) the population characteristics are identical versus 2) they are different. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical. The Census Bureau uses 90 percent confidence level for hypothesis testing but stricter confidence levels – for example 95 percent or higher – may also be used.

To perform the most common test, compute the difference  $X_A - X_B$ , where  $X_A$  and  $X_B$  are sample estimates of the characteristics of interest. A later section explains how to derive an

estimate of the standard error of the difference  $X_A - X_B$ . Let that standard error be  $S_{DIFF}$ . Calculate a test statistic  $|z|$ , as  $|z| = |X_A - X_B|/S_{DIFF}$ . For large sample sizes, one can conclude the difference is significant at the 90 percent confidence level if  $|z|$  is larger than the critical value of 1.645.

Confidence intervals can also be used to test for significant difference between two sample estimates. If  $X_A - X_B$  is between  $(-1.645 \times S_{DIFF})$  and  $(+1.645 \times S_{DIFF})$ , no conclusion about the characteristics is justified at the 10 percent significance level. If, on the other hand  $X_A - X_B$ , is smaller than  $(-1.645 \times S_{DIFF})$  or larger than  $(+1.645 \times S_{DIFF})$ , the observed difference is significant at the 10 percent significance level. In this event, it is commonly accepted practice to say that the characteristics are significantly different. In accordance with the Census Bureau's Statistical Quality Standards, we recommend that users report only those differences that are significant at the 10 percent significance level or better. Of course, sometimes this conclusion will be wrong. When the characteristics are the same, there is a 10 percent chance of concluding that they are different.

Note that as more tests are performed, more erroneous significant differences will occur. For example, at the 10 percent significance level, if 100 independent hypothesis tests are performed in which there are no real differences, it is likely that about 10 erroneous differences will occur. Therefore, the significance of any single test should be interpreted cautiously. A Bonferroni correction can be used to account for this potential problem. The Bonferroni method involves dividing your stated level of significance by the number of tests you are performing (Sedgwick, 2014; Stoline, 1981). This correction results in a conservative test of significance.

Because of the large standard errors involved, there is little chance that estimates will reveal useful information when computed on a small number of sample cases. We recommend a minimum of 30 unweighted cases for computing both cross-sectional and longitudinal estimates. Nonsampling error in one or more of the small number of cases providing the estimation can also cause large relative error in that particular estimate. Care must be taken in the interpretation of small differences since even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test.

### 3.3 Calculating Standard Errors for SIPP Estimates

There are two main ways we calculate the Standard Errors (SEs) for SIPP Estimates. They are as follows:

- Direct estimates using replicate weight methods (highly recommended); and
- Generalized variance function parameters (denoted as  $a$  and  $b$ ).

Replicate weight methods provide the most accurate variance estimates but may require more computing resources and expertise on the part of the user. The Generalized Variance Function (GVF) parameters provide a method of balancing accuracy with resource usage as well as having a smoothing effect on SE estimates across time. SIPP uses the Replicate Weighting Method to produce GVF parameters (see Chapter 7 of K. Wolter, *Introduction to Variance Estimation*, for more information). Table 8 shows calendar year 2019 SIPP key estimates their standard errors, coefficient of variation and item response rates computed from CY20 data. Standard errors for the key estimates were computed using replicate weights and Fay's BRR method.

Table 8. Calendar year 2019 SIPP key estimates computed from CY20 data.

Key Estimate	Estimate	Standard Error	Coefficient of Variation	Item Response Rate
Median annual household earnings	45,110	610	1.35	68.03
Median annual household income	65,760	544	0.83	51.53
Median household net worth	118,000	2,073	1.76	39.46
Poverty rate	11.48	0.29	2.50	58.64
Percent of households receiving means-tested benefits <sup>1</sup>	20.72	0.52	2.49	79.80
Percent receiving means-tested transfer income <sup>2</sup>	12.79	0.26	2.01	84.87
Percent receiving Supplemental Nutrition Assistance Program (SNAP) benefits	9.91	0.25	2.52	92.76
Percent receiving Supplemental Security Income (SSI)	2.86	0.08	2.96	86.50
Percent receiving income from social insurance programs <sup>3</sup>	24.98	0.16	0.64	85.50
Percent receiving Old Age, Survivors and Disability Insurance (OASDI) income	22.32	0.14	0.64	87.17
Percent covered by public health insurance	37.32	0.29	0.77	87.08
Percent covered by Medicaid	20.74	0.28	1.34	88.49
Percent covered by Medicare	19.11	0.11	0.57	88.42
Percent covered by private health insurance	66.59	0.34	0.51	91.26
Percent covered by employer-sponsored health insurance	56.30	0.34	0.61	87.16
Percent uninsured	8.42	0.18	2.18	87.02
Percent with a retirement plan or account	43.77	0.28	0.64	86.60

Key Estimate	Estimate	Standard Error	Coefficient of Variation	Item Response Rate
Percent of workers working two or more jobs	17.45	0.29	1.67	89.36
Percent of workers who are self-employed	13.99	0.28	2.03	91.01
Percent of workers who worked full-time, year-round	61.26	0.33	0.53	82.92
Percent of workers who work from home at least one day per week	18.28	0.28	1.53	99.98
Children ever born per 1,000 women age 15 to 50	1,128	11.98	1.06	46.86
Children ever born per 1,000 men age 15 to 50	901	12.27	1.36	37.54
Percent of mothers with multi-partner fertility	18.42	0.37	1.99	67.55
Percent of fathers with multi-partner fertility	16.54	0.38	2.30	63.47
Percent with a disability (age 1+)	20.10	0.21	1.03	87.46
Percent of households who experienced a change in composition	15.78	0.31	1.96	100.00

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2020.

<sup>1</sup> Means-tested benefits include Medicaid, Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), General Assistance (GA), Temporary Assistance for Needy Families (TANF), or special Supplemental Nutrition Program for Women Infants and Children (WIC).

<sup>2</sup> Means-tested transfer income include SNAP, SSI, TANF, WIC, and GA.

<sup>3</sup> Social insurance programs include Old Age, Survivors and Disability Insurance (OASDI), Veteran Affairs (VA) benefits, Workers Compensation (WC), and Unemployment Compensation (UC).

### 3.4 Standard Error Parameters and Tables and Their Use

Most SIPP estimates have greater standard errors than those obtained through a simple random sample because of SIPP's two-stage clustered sample design. To derive standard errors that would be applicable to a wide variety of estimates and could be prepared at a moderate cost, a few approximations were required.

Estimates with similar standard error behavior were grouped together and two parameters (denoted as  $a$  and  $b$ ) were developed to approximate the standard error behavior of each group of estimates. Because the actual standard error behavior was not identical for all estimates within a group, the standard errors computed from these parameters provide an indication of the order of magnitude of the standard error for any specific estimate. These  $a$  and  $b$  parameters vary by characteristic and by demographic subgroup to which the estimate applies. Tables 9, 10, and 11 provide  $a$  and  $b$  parameters for the core domains to be used for the CY20 cross-sectional (monthly and calendar year) and longitudinal estimates.

The creation of appropriate  $a$  and  $b$  parameters for the types of estimates previously given in Section 1.5 are described below.

1. The number of people who have ever experienced a characteristic during a given time period.

The appropriate  $a$  and  $b$  parameters are obtained directly from Tables 9, 10, or 11. The choice of parameter depends on the weights used, the characteristic of interest, and the demographic subgroup of interest.

2. Amount of a characteristic accumulated by people during a given time period.  
The appropriate  $a$  and  $b$  parameters are also taken directly from Tables 9, 10, or 11.
3. The average number of consecutive months of possession of a characteristic per spell (i.e., the average spell length for a characteristic) during a given time period.

Start with the appropriate base  $a$  and  $b$  parameters from Tables 9, 10, or 11. The parameters are then inflated by an additional factor,  $g$ , to account for people who experience multiple spells during the time period of interest. This factor is computed by:

$$g = \frac{\sum_{i=1}^n m_i^2}{\sum_{i=1}^n m_i} \quad (2)$$

where there are  $n$  people with at least one spell and  $m_i$  is the number of spells experienced by person  $i$  during the time period of interest.

4. The number of month-to-month changes in the status of a characteristic (i.e., number of transitions) summed over every set of two consecutive months during the time period of interest.

Obtain a set of adjusted  $a$  and  $b$  parameters exactly as just described in 3, then multiply these parameters by an additional factor. Use 1.0 if the time period of interest is two months and 2.0 for a longer time period. (The factor of 2.0 is based on the conservative assumption that each spell produces two transitions within the time period of interest.)

5. Monthly estimates of a characteristic averaged over a number of consecutive months.  
Appropriate base  $a$  and  $b$  parameters are taken from Tables 9, 10, or 11.

### 3.5 Standard Errors of Estimated Numbers

The approximate standard error,  $s_x$ , of an estimated number of persons, households, families, unrelated individuals and so forth, can be obtained using formula 3. Note that this method should be applied to dollar values

$$s_x = \sqrt{ax^2 + bx} \quad (3)$$

Here  $x$  is the size of the estimate and  $a$  and  $b$  are the appropriate parameters from Tables 9, 10, or 11 associated with the characteristic being estimated and the appropriate reference period.

#### Illustration 1.

Suppose SIPP estimates based on CY20 data show that there were 6,371,000 females aged 25 to 44 with a monthly earned income greater than \$6,000 in September 2019. The appropriate parameters and factor from Table 9 are:

$$a = -0.00005487 \quad b = 7,483$$

Using Formula (3), the approximate standard error is:

$$s_x = \sqrt{(-0.00005487 \times 6,371,000^2) + (7,483 \times 6,371,000)} = 213,183 \text{ females.}$$

The approximate 90 percent confidence interval as shown by the data is from 6,020,314 to 6,721,686 females (*i. e.*,  $6,371,000 \pm 1.645 \times 213,183$ ). In other words, 90 percent of confidence intervals from all possible samples will contain the actual number of females aged 25 to 44 with a monthly earned income greater than \$6,000 in September 2019 in the population.

### 3.6 Standard Error of a Mean

A mean is defined here to be the average quantity of some item (other than persons, families, or households) per person, family or household. For example, it could be the average monthly household income for a specified demographic. The standard error of a mean can be approximated by Formula (4) below. Because of the approximations used in developing Formula (4), an estimate of the standard error of the mean obtained from this formula will generally underestimate the true standard error. The formula used to estimate the standard error of a mean  $\bar{x}$  is:

$$s_{\bar{x}} = \sqrt{\left(\frac{b}{y}\right) s^2}, \quad (4)$$

where  $y$  is the size of the base,  $s^2$  is the estimated population variance of the item and  $b$  is the parameter associated with the particular type of item.

The population variance  $s^2$  may be estimated by one of two methods. In both methods, we assume  $x_i$  is the value of the item for  $i^{th}$  unit. (A unit may be person, family, or household). To use the first method, the range of values for the item is divided into  $c$  intervals. The lower and upper boundaries of interval  $j$  are  $Z_{j-1}$  and  $Z_j$ , respectively. Each unit,  $x_i$ , is placed into one of  $c$  intervals such that  $Z_{j-1} < x_i \leq Z_j$ . The estimated population mean,  $\bar{x}$ , and variance,  $s^2$ , are given by the formulas:

$$\bar{x} = \sum_{j=1}^c p_j m_j \quad (5a)$$

$$s^2 = \sum_{j=1}^c p_j m_j^2 - \bar{x}^2 \quad (5b)$$

where  $m_j = (Z_{j-1} + Z_j)/2$ , and  $p_j$  is the estimated proportion of units in the interval  $j$ . The most representative value of the item in the interval  $j$  is assumed to be  $m_j$ . If the interval  $c$  is open-ended, or no upper interval boundary exists, then an approximate value for  $m_c$  is

$$m_c = \frac{3}{2} Z_{c-1}.$$

In the second method, the estimated population mean,  $\bar{x}$ , and variance,  $s^2$  are given by:

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (6a)$$

$$s^2 = \frac{\sum_{i=1}^n w_i x_i^2}{\sum_{i=1}^n w_i} - \bar{x}^2 \quad (6b)$$

where there are  $n$  units with the item of interest and  $w_i$  is the final weight for  $i^{th}$  unit. (Note that  $\sum w_i = y$ .)

## Illustration 2.

### Method 1

Suppose that based on CY20 data, the distribution of annual income for persons aged 25 to 34 who were employed for all 12 months of 2019 is given in Table A.

Table A. Hypothetical Distribution of Annual Cash Income among People 25 to 34 Years Old (Not Actual Data, Only Use for Calculation Illustrations)

Interval of Annual Cash Income	Number of People in Each Interval (in thousands)	Cumulative Number of People with at Least as Much as Lower Bound of Each Interval (in thousands)	Percent of People with at Least as Much as Lower Bound of Each Interval
under \$5,000	370	23,527	100
\$5,000 to \$7,500	302	23,158	98.4
\$7,500 to \$9,999	447	22,856	97.1
\$10,000 to \$12,499	685	22,409	95.2
\$12,500 to \$14,999	935	21,724	92.3
\$15,000 to \$17,499	1,113	20,789	88.4
\$17,500 to \$19,999	1,298	19,675	83.6
\$20,000 to \$29,999	5,496	18,377	78.1
\$30,000 to \$39,999	4,596	12,881	54.7
\$40,000 to \$49,999	3,121	8,285	35.2
\$50,000 to \$59,999	1,902	5,164	21.9
\$60,000 to \$69,999	1,124	3,262	13.9
\$70,000 and above	2,138	2,138	9.1

Using these data, the mean monthly cash income for persons aged 25 to 34 is \$38,703.

Applying Formula (5b), the approximate population variance,  $s^2$ , is:

$$s^2 = \left( \frac{370}{23,527} \right) (2,500)^2 + \dots + \left( \frac{2,138}{23,527} \right) (105,000)^2 - (38,703)^2 = 649,442,787$$

Using Formula (4) and a  $b$  parameter of 7,483 from Table 9, the estimated standard error of a mean  $\bar{x}$  is:

$$s_{\bar{x}} = \sqrt{\frac{7,483}{23,527,000} \times 649,442,787} = \$454$$

Thus, the approximate 90 percent confidence interval as shown by the data ranges from \$37,956 to \$39,450.



### Method 2

Suppose that we are interested in estimating the average length of spells of SNAP reciprocity from January 2017 to December 2019, for the Black subpopulation. Also, suppose there are only 10 sample people in the subpopulation who were recipients of SNAP benefits. This example is a hypothetical situation used for illustrative purposes only; actually, 10 sample cases would be too few for a reliable estimate. The number of consecutive months of SNAP reciprocity during the reference period of interest and three-year longitudinal weights (*FINYR3*) are given in Table B below for each sample person:

Table B. Hypothetical Supplemental Nutrition Assistance Program (SNAP) Reciprocity Spells from January 2017 to December 2019 among SIPP Sample Persons (Not Actual Data, Only Use for Calculation Illustrations).

Sample Person	Spell Length in Months	Three-Year Longitudinal Weight ( <i>FINYR3</i> )
1	4, 3	12,728
2	5	3,480
3	9	3,492
4	3, 3, 2	7,491
5	12	9,936
6	12	15,140
7	4, 1	17,589
8	7	8,448
9	6	27,195
10	4	12,626

Using Formula (6a), the average spell of SNAP reciprocity is estimated to be:

$$\bar{x} = \frac{(12,728)(4) + (12,728)(3) + \dots + (12,626)(4)}{12,728 + 12,728 + \dots + 12,626} = \frac{859,519}{163,424} = 5.26$$

The standard error will be computed by Formula (4). First, the estimated population variance can be obtained by Formula (6b):

$$s^2 = \frac{(12,728)(4)^2 + (12,728)(3)^2 + \dots + (12,626)(4)^2}{12,728 + 12,728 + \dots + 12,626} - 5.26^2 = 11.24(\text{months})^2$$

Next, the base *b* parameter of 24,300 is taken from Table 11 and multiplied by the factor computed from Formula (2):

$$g = \frac{2^2 + 1 + 1 + 3^2 + 1 + 1 + 2^2 + 1 + 1 + 1}{2 + 1 + 1 + 3 + 1 + 1 + 2 + 1 + 1 + 1} = 1.71$$

Therefore, the final  $b$  parameter is  $1.71 \times 24,300 = 41,553$  and the standard error of the mean from Formula (4) is:

$$s_{\bar{x}} = \sqrt{\frac{(41,553)(11.24)}{163,424}} = 1.69 \text{ months}$$

### 3.7 Standard Error of an Aggregate

An aggregate is defined to be the total quantity of an item summed over all the units in a group. The standard error of an aggregate can be approximated using Formula (7). As with the estimate of the standard error of a mean, the estimate of the standard error of an aggregate will generally underestimate the true standard error. Let  $y$  be the size of the base,  $s^2$  be the estimated population variance of the item obtained using Formula (5b) or Formula (6b) and  $b$  be the parameter associated with the particular type of item. The standard error of an aggregate is:

$$s_x = \sqrt{b \times y \times s^2}. \quad (7)$$

### 3.8 Standard Errors of Estimated Percentages

The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends upon both the size of the percentage and the size of the total upon which the percentage is based. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more. For example, the percent of people employed is more reliable than the estimated number of people employed. When the numerator and denominator of the percentage have different parameters, use the parameter (and appropriate factor) of the numerator. If proportions are presented instead of percentages, note that the standard error of a proportion is equal to the standard error of the corresponding percentage divided by 100.

There are two types of percentages commonly estimated. The first is the percentage of people sharing a particular characteristic such as the percent of people owning their own home. The second type is the percentage of money or some similar concept held by a particular group of people or held in a particular form. Examples are the percent of total wealth held by people with high income and the percent of total income received by people on welfare.

For the percentage of people, the approximate standard error,  $s_{(x,p)}$ , of the estimated percentage  $p$  can be obtained by the formula:

$$s_{(x,p)} = \sqrt{\frac{b}{x}(p)(100 - p)}, \quad (8)$$

Here  $x$  is the size of the subclass of social units which is the base of the percentage,  $p$  is the percentage ( $0 < p < 100$ ), and  $b$  is the parameter associated with the characteristic in the numerator.

### Illustration 3.

Suppose that using the *FINYR2* weights it was estimated that 15,120,000 males were in poverty in 2018 and an estimated 4.78 percent of them exited poverty in 2019. Using Formula (8), with a  $b$  parameter of 18,870 from Table 10, the approximate standard error is:

$$s_{(x,p)} = \sqrt{\frac{18,870}{15,120,000} \times 4.78 \times (100 - 4.78)} = 0.75 \text{ percent}$$

Consequently, the 90 percent confidence interval as shown by these data is from 3.55 percent to 6.01 percent.

For percentages of money, a more complicated formula is required. A percentage of money will usually be estimated in one of two ways. It may be the ratio of two aggregates:

$$p_R = 100 \left( \frac{x_A}{x_B} \right),$$

or it may be the ratio of two means with an adjustment for different bases:

$$p_R = 100 \left( \hat{p}_A \left( \frac{\bar{x}_A}{\bar{x}_B} \right) \right),$$

where  $x_A$  and  $x_B$  are aggregate money figures,  $\bar{x}_A$  and  $\bar{x}_B$  are mean money figures, and  $\hat{p}_A$  is the estimated number in group A divided by the estimated number in group B. In either case, we estimate the standard error as

$$s_{(x,p_R)} = \sqrt{\left( \frac{\hat{p}_A \bar{x}_A}{\bar{x}_B} \right)^2 \left[ \left( \frac{s_p}{\hat{p}_A} \right)^2 + \left( \frac{s_A}{\bar{x}_A} \right)^2 + \left( \frac{s_B}{\bar{x}_B} \right)^2 \right]}, \quad (9)$$

where  $s_p$  is the standard error of  $\hat{p}_A$ ,  $s_A$  is the standard error of  $\bar{x}_A$  and  $s_B$  is the standard error of  $\bar{x}_B$ . To calculate  $s_p$ , use Formula (8). The standard errors of  $\bar{x}_B$  and  $\bar{x}_A$  may be calculated using Formula (4).

It should be noted that there is frequently some correlation between  $\hat{p}_A$ ,  $\bar{x}_B$ , and  $\bar{x}_A$ . Depending on the magnitude and sign of the correlations, the standard error will be over or underestimated.

Illustration 4.

In 2019, 7.21 percent of households owned rental property. The mean value of rental property among these households who own them is \$454,100, the mean value of assets among all households is \$777,600 and the corresponding standard errors are 0.180 percent, \$17,077, and \$50,009 respectively. Then the percent of all household assets held in rental property is:

$$100 \left( 0.0721 \times \frac{454,100}{777,600} \right) = 4.21 \text{ percent}$$

Using Formula (9), the corresponding standard error is

$$\begin{aligned} s_{(x,p_R)} &= \sqrt{\left( \frac{0.0721 \times 454,100}{777,600} \right)^2 \left[ \left( \frac{0.00180}{0.0721} \right)^2 + \left( \frac{17,077}{454,100} \right)^2 + \left( \frac{50,009}{777,600} \right)^2 \right]} \\ &= 0.331 \text{ percent} \end{aligned}$$

**3.9 Standard Error of a Difference**

The standard error of a difference between two sample estimates is approximately equal to

$$s_{(x-y)} = \sqrt{s_x^2 + s_y^2 - r s_x s_y} \quad (10)$$

where  $s_x$  and  $s_y$  are the standard errors of the estimates  $x$  and  $y$ .

The estimates can be numbers, percent, ratios, etc. The correlation between  $x$  and  $y$  is represented by  $r$ . The above formula assumes that the correlation coefficient between the characteristics estimated by  $x$  and  $y$  is non-zero. If no correlations have been provided for a given set of  $x$  and  $y$  estimates, assume  $r = 0$ . However, if the correlation is positive (negative), then this assumption will tend to cause overestimates (underestimates) of the true standard error.

Illustration 5.

Supposed that SIPP estimates show 944,500 Hispanic males and 1,007,000 Hispanic females received Federal and/or State administered Supplemental Security Income (SSI) in December 2019. Then, using the *Male Hispanic* GVF parameters ( $a = -0.0002574$  and  $b = 7,846$ ) and the *Female Hispanic* GVF Parameters ( $a = -0.0002586$  and  $b = 7,846$ ) from Table 9 and Formula (3), the standard errors of these numbers are approximately 84,740 and 87,400 respectively. The difference in sample estimates is 62,500 and using Formula (10), the approximate standard error of the difference is:

$$\sqrt{84,740^2 + 87,400^2} = 121,736$$

Suppose that it is desired to test at the 10 percent significance level whether the number of female Hispanic SSI recipients was different from the number of male Hispanic SSI recipients in December 2019. To perform the test, compute the Z statistic and compare it to the critical value of 1.645 for the 10 percent significance level.

$$Z = \frac{62,500}{121,736} = 0.51$$

Since 0.51 is less than 1.645, we can conclude that the number of male and female Hispanic SSI recipients are not significantly different at the 10 percent significance level.

### 3.10 Standard Error of a Median

The median quantity of some item such as income for a given group of people is that quantity such that at most 50 percent of the group have less and at most 50 percent of the group have more. The sampling variability of an estimated median depends upon the form of the distribution of the item as well as the size of the group. To calculate standard errors on medians, the procedure described below may be used.

The median, like the mean, can be estimated using either data which have been grouped into intervals or ungrouped data. If grouped data are used, the median is estimated using Formulas (11) or (12) with  $q = 0.5$ , where  $q$  is the proportion of a group possessing characteristics of interest. If ungrouped data are used, the data records are ordered based on the value of the characteristic, then the estimated median is the value of the characteristic such that the weighted estimate of half – i.e. 50 percent – of the subpopulation falls at or below that value and half is at or above that value. Note that the method of standard error computation which is presented here requires the use of grouped data. Therefore, it should be easier to compute the median by grouping the data and using Formula (11) or (12).

An approximate method for measuring the reliability of an estimated median is to determine a confidence interval about it. (See the section on sampling variability for a general discussion of confidence intervals.) The following procedure may be used to estimate the 68 percent confidence limits and hence the standard error of a median based on sample data.

1. Determine, using  $p = 50$  in Formula (8), the standard error of an estimate of 50 percent of the group.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the item within the group, calculate the quantity of the item such that the percent of the group with more of the item is equal to the smaller percentage found in step 2. This quantity will be the upper limit for the 68 percent confidence interval. In a similar fashion, calculate the quantity of the item such that the

percent of the group with more of the item is equal to the larger percentage found in step 2. This quantity will be the lower limit for the 68 percent confidence interval.

4. Divide the difference between the two quantities determined in step 3 by two to obtain the standard error of the median.

To perform step 3, it will be necessary to interpolate. Different methods of interpolation may be used. The most common are simple linear interpolation and Pareto interpolation. The appropriateness of the method depends on the form of the distribution around the median. If density is declining in the area, then we recommend Pareto interpolation. If density is fairly constant in the area, then we recommend linear interpolation. Note, however, that Pareto interpolation can never be used if the interval contains zero or negative measures of the item of interest. Interpolation is used as follows.

The quantity of the item such that  $p$  percent have more of the item is:

$$X_{pN} = A_1 \times \exp \left[ \left( \frac{\ln \left( \frac{qN}{N_1} \right)}{\ln \left( \frac{N_2}{N_1} \right)} \right) \ln \left( \frac{A_2}{A_1} \right) \right] \quad (11)$$

if Pareto Interpolation is indicated and:

$$X_{pN} = \left[ A_1 + \left( \frac{qN - N_1}{N_2 - N_1} \right) (A_2 - A_1) \right], \quad (12)$$

if linear interpolation is indicated, where:

$N$	is the size of the group,
$A_1$ and $A_2$	are the lower and upper bounds, respectively, of the interval in which $X_{pN}$ falls
$N_1$ and $N_2$	are the estimated number of group members owning more than $A_1$ and $A_2$ , respectively
$exp$	refers to the exponential function and
$ln$	refers to the natural logarithm function
$q$	is the corresponding proportion for $p$ i.e., $q = p/100$

#### Illustration 6.

To illustrate the calculations for the sampling error on a median, we return to Table A from illustration 2. The median annual income for this group using Formula (11) is \$31,828. The size of the group is 23,527,000.

1. Using Formula (8), the standard error of 50 percent on a base of 23,527,000 is about 0.89 percentage points.
2. Following step 2, the two percentages of interest are 49.11 and 50.89.
3. By examining Table A, we see that the percentage 49.11 falls in the income interval from \$30,000 to \$39,999. (Since 54.7 percent receive more than \$30,000 per annum, the dollar value corresponding to 49.11 must be between \$30,000 and \$40,000.) Thus,  $A_1 = \$30,000$ ,  $A_2 = \$40,000$ ,  $N_1 = 12,881,000$  and  $N_2 = 8,285,000$ .

In this case, we decided to use Pareto interpolation. Therefore, using Formula (11), the upper bound of a 68 percent confidence interval for the median is

$$\$30,000 \times \exp \left[ \left( \frac{\ln \left( \frac{0.4911 \times 23,527,000}{12,881,000} \right)}{\ln \left( \frac{8,285,000}{12,881,000} \right)} \right) \times \ln \left( \frac{40,000}{30,000} \right) \right] = \$ 32,203 .$$

Also, by examining Table A, we see that 50.89 falls in the same income interval. Thus,  $A_1, A_2, N_1$  and  $N_2$  are the same. We also use Pareto interpolation for this case. So, the lower bound of a 68 percent confidence interval for the median is

$$\$30,000 \times \exp \left[ \left( \frac{\ln \left( \frac{0.5089 \times 23,527,000}{12,881,000} \right)}{\ln \left( \frac{8,285,000}{12,881,000} \right)} \right) \times \ln \left( \frac{40,000}{30,000} \right) \right] = \$ 31,464 .$$

Thus, the 68 percent confidence interval on the estimated median is from \$31,464 to \$32,203.

4. Then the approximate standard error of the median is

$$\frac{\$ 32,203 - \$ 31,464}{2} = \$369.5$$

### 3.11 Standard Errors of Ratios of Means and Medians

The standard error for a ratio of means or medians is approximated by:

$$s_{\frac{x}{y}} = \sqrt{\left(\frac{x}{y}\right)^2 \left[ \left(\frac{s_y}{y}\right)^2 + \left(\frac{s_x}{x}\right)^2 \right]}, \quad (13)$$

where  $x$  and  $y$  are the means or medians, and  $s_x$  and  $s_y$  are their associated standard errors. Formula (13) assumes that the means are not correlated. If the correlation between the population means estimated by  $x$  and  $y$  are actually positive (negative), then this procedure

will tend to produce overestimates (underestimates) of the true standard error for the ratio of means.

### 3.12 Standard Errors Using Software Packages

Standard errors and their associated variance, calculated by statistical software packages such as SAS or Stata, do not accurately reflect the SIPP's complex sample design. Erroneous conclusions will result if these standard errors are used directly. We provide adjustment factors by characteristics that should be used to correctly compensate for likely under-estimates. The factors called design effects (DEFF), available in Tables 9 through 11, must be applied to SAS or Stata generated variances. The square root of DEFF can be directly applied to similarly generated standard errors. These factors approximate design effects which adjust statistical measures for sample designs more complex than simple random sample.

Replicate weights for SIPP are also provided and can be used to estimate more accurate standard errors and variances. While replicate weighting methods require more computing resources, many statistical software packages, including SAS, have procedures that simplify the use of replicate weights for users. To calculate variances using replicate weights use the formula:

$$Var(\theta_0) = \frac{1}{G(0.5)^2} \times \sum_{i=1}^G (\theta_i - \theta_0)^2 \quad (14)$$

where  $G=240$  is the number of replicates,  $\theta_0$  is the estimate using full sample weights, and  $\theta_i$  is the estimate using the  $i^{th}$  replicate weight. Replicate weights are created using Fay's method, with a Fay coefficient of 0.5 (Chakrabarty, 1993; Fay, 1984).

Instead of direct computation, various SAS procedures include options to use replicate weights when estimating standard errors or variances. Fay's method is a variation of the balanced repeated replication (BRR) method, so within SAS SURVEY procedures include the `VARMETHOD=BRR(FAY=0.5)` option in the PROC statement and specify the replicate weights with a `REPWEIGHTS` statement. Other computer packages have similar methods.

Formula (14) produces variance estimates close to zero for the median when multiple observations have value equal to the median. In this case, two methods can be used to estimate the variance of the median. The first technique incorporates replicate weights in Woodruff's method for estimating variability (Woodruff, 1952). Gossett et al. (2002) documents the procedure for combining Woodruff's method with Jackknife replication and provides sample codes adapted by Mack and Tekansik (2011) for Fay's BRR. The second method uses `VARMETHOD=TAYLOR` option, a direct application of Woodruff's method, along with the cluster (GHLFSAM variable) and strata statements (GVARSTR variable) instead of replicate weights to account for SIPP's complex design.



### Illustration 7.

In SAS, there are a set of SURVEY procedures that can use the replicate weights to estimate standard errors. All of these SAS procedures use the same VARMETHOD = BRR(FAY=0.5) option and REPWEIGHTS statement. The SURVEYMEANS procedure is used to estimate statistics such as means, totals, proportions, quantiles, and ratios for a survey sample, SURVEYFREQ is for frequency tables and cross-tabulations, SURVEYREG is for regression analysis, SURVEYLOGISTIC is for logistic regression analysis, and SURVEYPHREG is for proportional hazards regression analysis.

An example syntax for estimating the mean and median of the total household income (*THTOTINC*) for any month in 2019 using SIPP replicate weights is:

```
proc surveymeans data=pu202013 varmethod=brr(Fay=0.5) mean median14;
  var THTOTINC;
  weight WPFINWGT;
  repweights REPWGT1-REPWGT240;
run;
```

Alternatively, to fit a regression model with the total household income (*THTOTINC*) as the response variable and several household characteristics such as age (*TAGE*), race (*ERACE*), Hispanic origin (*EORIGIN*), and educational attainment (*EEDUC*) of the reference person, number of persons in the household (*RHNUMPER*), tenure (*EEHC\_TEN*), and region (*TEHC\_REGION*) as the predictor variables the SAS syntax is:

```
proc surveyreg data=pu2020 varmethod=brr(Fay=0.5);
  class ERACE EORIGIN EEHC_TEN TEHC_REGION;
  model THTOTINC = TAGE ERACE EORIGIN EEDUC RHNUMPER EEHC_TEN TEHC_REGION;
  weight WPFINWGT;
  repweights REPWGT1-REPWGT240;
run;
```

In Stata, the SVY command is used to fit a statistical model to a complex survey dataset. SVYSET is used to determine the survey design and provide information about the variance estimation. The following Stata syntax is equivalent to using SURVEYMEANS in SAS:

```
use pu2020.dta
svyset [pweight= WPFINWGT], brrweight(REPWGT1-REPWGT240) fay(.5) vce(brr) mse
svy: mean THTOTINC
```

<sup>13</sup> This snippet of code requires the analytic, final weight, and replicate weights variables – *THTOTINC*, *WPFINWGT*, and *REPWGT1-REPWGT240* respectively – are included in the dataset *pu2020* and that *pu2020* is restricted to the month of interest.

<sup>14</sup> The documentation for the Surveymeans procedure provides a list of “statistic keywords” that can be supplied to compute estimates of different statistics.

#### 4. References

- Chakrabarty, R. P. (1993). Variance estimation by users of SIPP micro-data files [Working Paper]. <https://www.census.gov/content/dam/Census/library/working-papers/1993/adrm/rr93-6.pdf>
- Fay, R. E. (1984). Some properties of estimates of variance based on replication methods. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 495-500.
- Gossett, J. M., Simpson, P., Parker, J.G., & Simon, W.L. (2002). How complex can complex survey analysis be with SAS? Proceedings of the 27th Annual SAS Users Group International Conference. <https://support.sas.com/resources/papers/proceedings/proceedings/sugi27/p266-27.pdf>
- Kalton, G. (1998). SIPP quality profile [Working Paper]. U.S. Census Bureau. <https://www.census.gov/content/dam/Census/library/working-papers/1998/demo/SEHSD-WP1998-11.pdf>
- Mack, S. & Tekansik, S. (2011). "Calculating Standard Errors and Confidence Intervals of Medians." U.S. Census Bureau Memorandum for Documentation dated May 6, 2011.
- Sedgwick, P. (2014). Multiple hypothesis testing and Bonferroni's correction. *BMJ: British Medical Journal*, 349. <https://www.ijstor.org/stable/26517649>
- Stoline, M. (1981). The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs. *The American Statistician*, 35(3), 134-141. <https://www.ijstor.org/stable/2683979>
- Tersine Jr., A.G. (2021a). "Survey of Income and Program Participation Cross-sectional Weighting Specifications for Calendar Year 2019 and Subsequent Years (WGT-34)." U.S. Census Bureau Memorandum from Anthony G. Tersine, Jr. to Jeffery Sission dated September 29, 2021.
- Tersine Jr., A.G. (2021b). "Survey of Income and Program Participation Longitudinal Weighting Specifications for Calendar Year 2019 and Subsequent Years (WGT-35)." U.S. Census Bureau Memorandum from Anthony G. Tersine, Jr. to Jeffery Sission. Not yet finalized
- U.S. Census Bureau. (2019). Survey of Income and Program Participation: 2014 Panel Users' Guide. Washington, D.C. <https://www.census.gov/programs-surveys/sipp/guidance/users-guide.html>

- U.S. Census Bureau. (2020). Survey of Income and Program Participation: 2018 Panel Users' Guide. Washington, D.C. Revised August 2021. <https://www.census.gov/programs-surveys/sipp/guidance/users-guide.html>
- U.S. Census Bureau. (2021). Survey of Income and Program Participation: 2020 Panel Users' Guide. Washington, D.C. <https://www.census.gov/programs-surveys/sipp/guidance/users-guide.html>
- Wolter, K. M. (2007). Chapter 7: Generalized variance functions, Introduction to Variance Estimation, 2nd Ed. New York: Springer, pp. 272-297.
- Woodruff, R. S. (1952). Confidence intervals for medians and other position measures. Journal of the American Statistical Association, 47(260), 635-646.

## 5. Generalized Variance Parameters and Tables

Table 9. Cross-Sectional Generalized Variance Parameters for CY20 data

Domain	Parameters		Design Effect <sup>15</sup>
	<i>a</i>	<i>b</i>	
<b>Poverty and Program Participation, Persons 15+</b>			
Total	-0.00003013	7,965	2.943
Male	-0.00006224	7,965	
Female	-0.00005841	7,965	
<b>Income and Labor Force Participation, Persons 15+</b>			
Total	-0.00002831	7,483	2.765
Male	-0.00005848	7,483	
Female	-0.00005487	7,483	
<b>Other, Persons 0+</b>			
Total (or White)	-0.00002363	7,681	2.839
Male	-0.00004832	7,681	
Female	-0.00004626	7,681	
<b>Black, Persons 0+</b>			
Male	-0.00021150	9,100	3.363
Female	-0.00045140	9,100	
	-0.00039810	9,100	
<b>Hispanic, Persons 0+</b>			
Male	-0.00012900	7,846	2.899
Female	-0.00025740	7,846	
	-0.00025860	7,846	
<b>Households</b>			
Total (or White)	-0.00004994	6,552	2.421
Black	-0.00038140	6,552	
Hispanic	-0.00035080	6,552	

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2020.

### Notes on Domain Usage for Table 9

Poverty and Program Participation	Use these parameters for estimates concerning poverty rates, welfare program participation (e.g. Supplemental Security Income, SSI), and other programs for adults with low incomes.
Income and Labor Force	These parameters are for estimates concerning income, sources of income, labor force participation, economic well-being other than poverty, employment related estimates (e.g. occupation, hours worked a week), and other income, job, or employment related estimates.
Other Persons	Use the "Other Persons" parameters for estimates of total (or white) persons aged 0+ in the labor force, and all other characteristics not specified in this table, for the total or white population.
Black/Hispanic Persons	Use these parameters for estimates of Black and Hispanic persons 0+.
Households	Use these parameters for all household level estimates.

<sup>15</sup> Design Effect=b/sample interval where sample interval=2,706

Table 10. Longitudinal Generalized Variance Parameters for CY20 FINYR2

Domain	Parameters		Design Effect <sup>16</sup>
	<i>a</i>	<i>b</i>	
<b>Poverty and Program Participation, Persons 15+</b>			
Total	-0.00007140	18,870	6.973
Male	-0.00014750	18,870	
Female	-0.00013840	18,870	
<b>Income and Labor Force Participation, Persons 15+</b>			
Total	-0.00006368	16,830	6.220
Male	-0.00013150	16,830	
Female	-0.00012340	16,830	
<b>Other, Persons 0+</b>			
Total (or White)	-0.00005484	17,820	6.585
Male	-0.00011210	17,820	
Female	-0.00010730	17,820	
<b>Black, Persons 0+</b>			
Male	-0.00053970	23,220	8.581
Male	-0.00115200	23,220	
Female	-0.00101600	23,220	
<b>Hispanic, Persons 0+</b>			
Male	-0.00033730	20,520	7.583
Male	-0.00067310	20,520	
Female	-0.00067630	20,520	
<b>Households</b>			
Total (or White)	-0.00011090	14,760	5.455
Black	-0.00084010	14,760	
Hispanic	-0.00078650	14,760	

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2020.

#### Notes on Domain Usage for Table 10

Poverty and Program Participation	Use these parameters for estimates concerning poverty rates, welfare program participation (e.g. Supplemental Security Income, SSI), and other programs for adults with low incomes.
Income and Labor Force	These parameters are for estimates concerning income, sources of income, labor force participation, economic well-being other than poverty, employment related estimates (e.g. occupation, hours worked a week), and other income, job, or employment related estimates.
Other Persons	Use the "Other Persons" parameters for estimates of total (or white) persons aged 0+ in the labor force, and all other characteristics not specified in this table, for the total or white population.
Black/Hispanic Persons	Use these parameters for estimates of Black and Hispanic persons 0+.
Households	Use these parameters for all household level estimates.

<sup>16</sup> Design Effect=b/sample interval where sample interval=2,706

Table 11. Longitudinal Generalized Variance Parameters for CY20 FINYR3

Domain	Parameters		Design Effect <sup>17</sup>
	<i>a</i>	<i>b</i>	
<b>Poverty and Program Participation, Persons 15+</b>			
Total	-0.00007328	19,370	7.158
Male	-0.00015140	19,370	
Female	-0.00014210	19,370	
<b>Income and Labor Force Participation, Persons 15+</b>			
Total	-0.00006514	17,220	6.364
Male	-0.00013460	17,220	
Female	-0.00012630	17,220	
<b>Other, Persons 0+</b>			
Total (or White)	-0.00005595	18,180	6.718
Male	-0.00011440	18,180	
Female	-0.00010950	18,180	
<b>Black, Persons 0+</b>			
Male	-0.00056480	24,300	8.980
Female	-0.00120500	24,300	
	-0.00106300	24,300	
<b>Hispanic, Persons 0+</b>			
Male	-0.00034750	21,130	7.809
Female	-0.00069330	21,130	
	-0.00069660	21,130	
<b>Households</b>			
Total (or White)	-0.00011240	15,150	5.599
Black	-0.00084810	15,150	
Hispanic	-0.00080210	15,150	

Source: U.S. Census Bureau, Survey of Income and Program Participation, Calendar Year 2020.

#### Notes on Domain Usage for Table 11

Poverty and Program Participation	Use these parameters for estimates concerning poverty rates, welfare program participation (e.g. Supplemental Security Income, SSI), and other programs for adults with low incomes.
Income and Labor Force	These parameters are for estimates concerning income, sources of income, labor force participation, economic well-being other than poverty, employment related estimates (e.g. occupation, hours worked a week), and other income, job, or employment related estimates.
Other Persons	Use the "Other Persons" parameters for estimates of total (or white) persons aged 0+ in the labor force, and all other characteristics not specified in this table, for the total or white population.
Black/Hispanic Persons	Use these parameters for estimates of Black and Hispanic persons 0+.
Households	Use these parameters for all household level estimates.

<sup>17</sup> Design Effect=b/sample interval where sample interval=2,706